9H

УДК 004.934

А.В. Ниценко[1], Т.С. Хашан[2]
[1]Институт проблем искусственного интеллекта МОН и НАН Украины, г. Донецк, Украина
[2]Институт прикладной математики и механики НАН Украины, г. Донецк, Украина
khashan@mail.ru

Применение комплексного непрерывного вейвлет-преобразования Морле в обработке аудиосигналов

В статье предложен модифицированный метод фазового вокодера на основе комплексного непрерывного преобразования Морле. Комбинация указанного метода, линейной интерполяции и КИХ-фильтра позволила разработать алгоритм изменения темпа сигнала без изменения его тона.

В современных приложениях профессиональной обработки аудиоданных широко используются различные методы преобразований сигнала с целью создания эффектов высокого качества. Например, для сокращения времени прослушивания аудиосигнала, подготовки цифровых фонограмм, реставрации записей, синтезирования электронной музыки, для систем управления партитурой в реальном времени применяют методы изменения высоты тона и темпа сигнала. Они позволяют значительно сэкономить время, поскольку дают возможность исправлять ошибки или корректировать запись без выполнения новой.

Разработка качественных методов изменения темпа проигрывания сигнала без изменения его тона является актуальной. На текущий момент имеется несколько подходов к решению этой задачи.

Методы изменения темпа сигнала без изменения его тона

Алгоритмы SOLA и PSOLA [1] основаны на методах анализа и обработки сигнала во временной области. Недостатком данных методов является плохое качество звучания измененного сигнала и появление в нем различных артефактов.

Применение метода фазового вокодера [2] и ресемплирования сигнала позволяет частично устранить перечисленные недостатки. Ресемплирование сигнала меняет скорость воспроизведения сигнала за счет изменения его частоты дискретизации. Для этого обычно применяют методы интерполяции и фильтрации сигналов. Фазовый вокодер – это метод получения частотно-временнóго представления сигнала при помощи быстрого преобразования Фурье (БПФ). Однако БПФ не сохраняет временную информацию о сигнале, о времени появления той или иной частоты. Метод кратковременного преобразования Фурье (КВПФ) является серией оконных БПФ, перекрывающихся по времени и позволяющих за счет фреймового подхода сохранять временнýю информацию о спектре сигнала.

Один из традиционных подходов к смене высоты тона без изменения длительности сигнала состоит в вычислении двухмерного частотно-временного представления сигнала описанным выше методом. После этого интервал сетки анализа и фазы коэффициентов масштабируются для создания сетки синтеза. Далее выполняется обратное преобразование, восстанавливающее звуковой сигнал. Таким образом, алгоритм фазового вокодера можно описать последовательностью действий:

1. Вычисление фазового вокодера.
2. Преобразование коэффициентов в полярную форму.
3. Развертывание фазового спектра и умножение на коэффициент масштабирования c.
4. Восстановление сигнала по новой сетке синтеза.

Сетка анализа фазового вокодера очень похожа на частотно-временное представление сигнала, получаемое в результате непрерывного вейвлет-преобразования (НВП) [3-7]. Как и фазовый вокодер, НВП преобразует сигнал в его частотно-временное представление. Изменяя компоненты сетки анализа НВП, можно получить эффекты, подобные тем, что получаются при помощи вокодера. В отличие от КВПФ, в области высоких частот НВП обеспечивает лучшее временное разрешение, а в области низких частот – лучшее частотное разрешение. Главное преимущество вейвлетов – возможность полного восстановления сигнала, которую не обеспечивает Фурье-преобразование, в частности, из-за эффекта Гиббса [5]. Использование окон, которые борются с этим эффектом, ухудшает восстановление сигнала на участках его быстрых изменений.

Свойство переменного разрешения хорошо подходит для представления музыкальных и речевых сигналов и поэтому, по мнению авторов, применение НВП позволит получить лучший результат (т.е. улучшить качество звучания измененного сигнала), по сравнению с фазовым вокодером, у которого разрешение одинаково во всех частотных областях.

Целью данной работы является разработка методики и алгоритма обработки аудиосигнала на основе методов вейвлет-анализа применительно к решению задачи изменения темпа сигнала без изменения его тона.

Постановка задачи

Для достижения поставленной цели исследования необходимо:
– выбрать оптимальный базис вейвлет-разложения;
– на основе его частотных характеристик и правил вычисления вейвлет-преобразования разработать методику изменения тона сигнала;
– для изменения темпа сигнала и сохранения качества звучания выбрать оптимальный метод его ресэмплирования.

Выбор базиса вейвлета и его характеристик

Математический аппарат непрерывных вейвлет-преобразований (НПВ) хорошо описан в [3-5]. Качество анализа сигналов с помощью вейвлет-преобразований зависит от выбора информативного базиса.

В работах [6], [7] показано, что для исследования речевых (РС) и музыкальных сигналов наиболее информативным является непрерывный вейвлет Морле, т.к. его базис лучше аппроксимирует сигнал:

$$\psi(t) = e^{j5t} e^{-\frac{t^2}{2}} \quad (1)$$

Зная центр $\langle\omega\rangle$ и радиус Δ_Ψ частотного окна базового вейвлета ($\langle\omega\rangle = 5$, $\Delta_\Psi = 0{,}7071$ [6]), пользуясь правилом квантования масштабирующей переменной a и формулой для определения ширины частотного окна (2), можно получить значения границ полос пропускания вейвлет-фильтров на каждом уровне разложения m:

$$win_m = [\langle\omega\rangle / a_0^m - \Delta_\Psi / a_0^m ; \langle\omega\rangle / a_0^m + \Delta_\Psi / a_0^m],\ j_{\min} \le m \le j_{\max}. \quad (2)$$

Форма АЧХ зависит как от самого анализирующего вейвлета, так и от набора масштабов, на которых выполняется преобразование. Для этого вычисляются начальный (3) и конечный (4) уровни разложения, на которых осуществляется обработка и анализ сигнала:

$$j_{\min} = \operatorname{int}\left(\log_{a_0} \frac{f_d \langle \omega \rangle}{2\pi f_2}\right), \quad (3)$$

$$j_{\max} = \operatorname{int}\left(\log_{a_0} \frac{f_d \langle \omega \rangle}{2\pi f_1}\right), \quad (4)$$

где int() – округление до ближайшего целого.

В данной работе границы частотного диапазона исследуемых сигналов [f_1; f_2] и их частота дискретизации f_d обозначены следующими величинами: $f_1 = 50$Гц; $f_2 = 20$ КГц (у музыкального сигнала более широкий спектр, чем у РС); $f_d = 44100$Гц.

Начальный и конечный уровни разложения сигналов (шаг масштабирования $a_0 = 1{,}1$) определяются по формулам (3 – 4) и составляют значения $j_{min} = 6$ и $j_{max} = 70$.

Для изменения тона сигнала и соответственно для расчета фазового спектра сигнала необходимо применить комплексное непрерывное вейвлет-преобразование (КНВП) Морле.

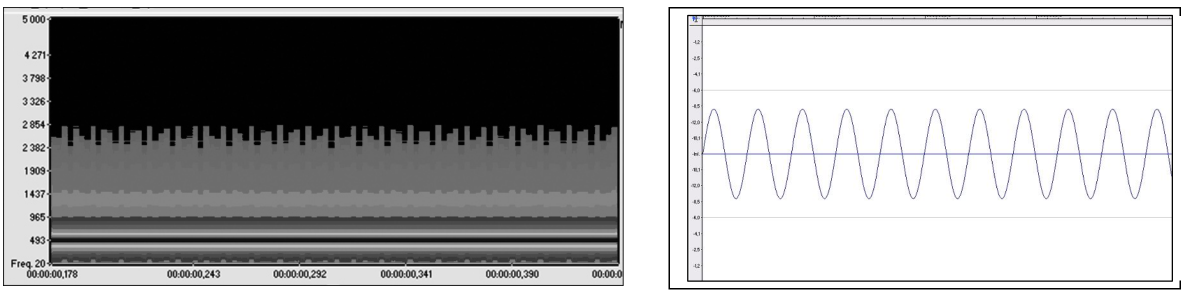
Методика применения КНВП Морле для построения фазового вокодера

1. Вычисление коэффициентов КНВП Морле:

$$d_{mn} = CDWT_f(m,n) = \sum_{i=0}^{N-1} f(i)\psi_{mn}(i)\Delta t \quad (5)$$

$m = \overline{1, P}$, $n = \overline{0, N-1}$, где Δt – шаг квантования по времени.

На рис. 1 показан пример спектра (а) исходного сигнала и его амплитудно-временное представление (б). Для наглядности примера был выбран монохроматический сигнал с частотой 500 ГЦ.



а) б)

Рисунок 1 – Исходный сигнал $w = 500\,Гц$: а) – спектр сигнала; б) амплитудно-временное представление сигнала

2. Вычисление амплитуды коэффициентов:

$$A(m,n) = \sqrt{CDWT_{re}(m,n)^2 + CDWT_{im}(m,n)^2}\,. \quad (6)$$

3. Вычисление фазового спектра

$$\varphi(m,n) = \operatorname{arctg}\left(CDWT_{im}(m,n) / CDWT_{re}(m,n)\right). \quad (7)$$

4. Развертывание фазового спектра путем прибавления к выборкам главного значения фазы (п. 3), величин, равных $2\pi m$, где m – целое число. Значения m определяются путем сравнения соседних отсчетов фазы, образовывая так называемую корректирующую последовательность.

5. Вычисление новых коэффициентов на каждом уровне разложения с учетом коэффициента сдвига c:

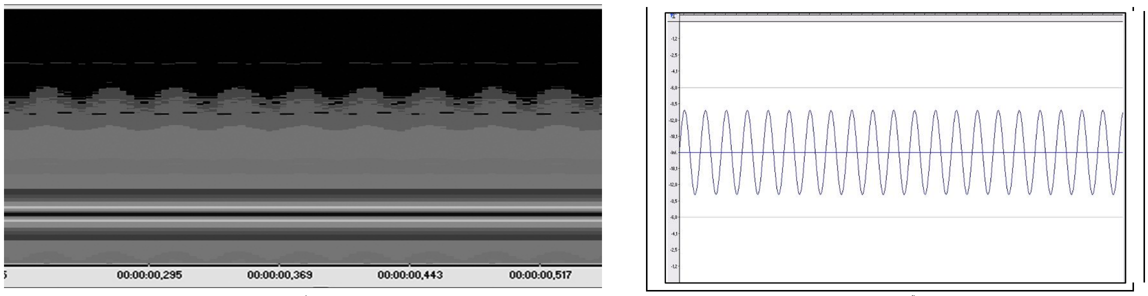
$$CDWT_{re}(m,n) = A(m,n)\cos(\varphi_{\text{unw}}(m,n)\cdot c) \quad (8)$$

$$CDWT_{im}(m,n) = A(m,n)\sin(\varphi_{\text{unw}}(m,n)\cdot c). \quad (9)$$

6. Сдвиг уровней на величину $\log_a c$.

7. Восстановление сигнала при помощи обратного НВП (рис. 2):

$$\tilde{f}(n) = \frac{E_f}{E_W}\sum_m \sum_{l=0}^{N-1} d_{ml}\psi_{ml}(n) \quad (10)$$



а) б)

Рисунок 2 – Восстановленный сигнал со сдвигом на октаву вверх ($w = 1000\,Гц$):
а) спектр сигнала; б) амплитудно-временное представление сигнала

Если необходимо выполнить изменение темпа без изменения тональности, то дополнительно выполняется ресемплирование – изменение частоты дискретизации в c раз.

Ресемплирование сигнала

Ресемплирование сигнала применяют для изменения темпа его воспроизведения в некоторое количество раз c с сохранением тональности звучания. Для этого изменяется высота тона в $\frac{1}{c}$ раз, а затем и частота дискретизации сигнала также в $\frac{1}{c}$ раз (то есть ресемплирование). Если $c < 1$, т.е. темп воспроизведения замедляется, то после изменения высоты тона в сторону повышения частота дискретизации увеличивается за счет добавления нужного количества новых отсчетов $f(x_c)$ между двумя уже существующими отсчетами сигнала $f(x_a)$ и $f(x_b)$ через равные промежутки времени. При этом значения новых отсчетов рассчитываются при помощи линейной интерполяции:

$$f(x_c) = f(x_a) + \frac{f(x_b) - f(x_a)}{(x_b - x_a)}(x_c - x_a). \quad (11)$$

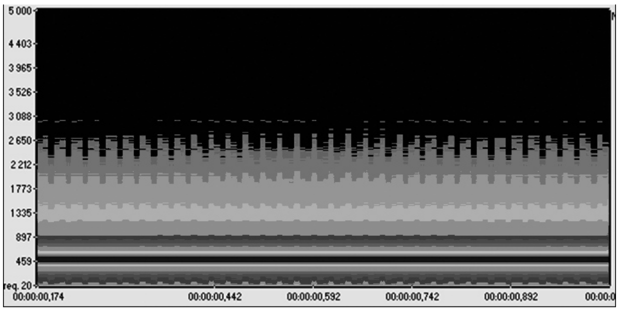
Количество вставляемых отсчетов равно $\frac{1}{c} - 1$.

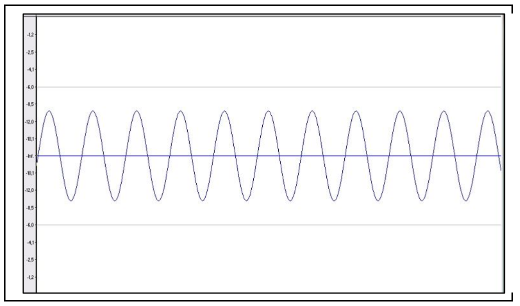
Для устранения высокочастотного шума выполняется сглаживание сигнала КИХ-фильтром.

Если $c > 1$ (темп увеличивается), то выполняется сдвиг тональности исходного сигнала в сторону низких частот, далее частота дискретизации уменьшается в c раз путем удаления через равные промежутки времени $c - 1$ отсчетов (прореживания) полученного сигнала.

Таким образом удается сохранить тональность звучания исходного сигнала.

На рис. 3 в качестве примера приведен ресемплированный сигнал, уменьшающий темп произведения ($c = 2$).





а) б)

Рисунок 3 – Ресемплированный сигнал: а) спектр сигнала;
б) амплитудно-временное представление сигнала

Выводы

В статье предложен алгоритм расчета вейвлет-коэффициентов КНВП Морле, учитывающий локализацию базисной вейвлет-функции на каждом уровне разложения; рассчитаны минимальный и максимальный уровни разложения по исследуемому базису, центральные частоты и полосы пропускания фильтров и их АЧХ; предложен модифицированный алгоритм фазового вокодера на основе НКВП Морле.

Разработан алгоритм изменения темпа сигнала без изменения его тона. Для этого применены линейная интерполяция и КИХ-фильтр.

Литература

1. Zolzer U. DAFX Digital Audio Effects. – West Sussex, England: Wiley, 2002. – P. 201-282.
2. Laroche Jean, Dolson Mark. Improved Phase Vocoder Time-Scale Modification of Audio IEEE Transactions on Speech and Audio Processing. – 1999. – Vol. 7, № 3.
3. Астафьева Н.М. Вейвлет-анализ: основы теории и некоторые приложения // Успехи физических наук. – 1998. – №11. – С. 1145-1170.
4. Поршнев С.В. Применение непрерывного вейвлет-преобразования для обработки широкополосных частотно-модулированных сигналов // Вычислительные методы и программирование. – 2003. – Т. 4. – С. 104-116.
5. Воробьев В.И., Грибунин В.Г. Теория и практика вейвлет-преобразования. – СПб.: ВУС, 1999. – 208 с.
6. Ермоленко Т.В. Применение вейвлет-анализа для предварительной обработки речевых голосовых сигналов и пофонемного распознавания изолированных слов: Дисс... канд. Техн. наук: Донецк, 2008. – 172 с.
7. Peter De Gersem, Bart De Moor, Marc Moonen. Applications of the Continuous Wavelet Transform in the Processing of Musical Signals // 13th International Conference on Digital Signal Processing (DSP97). – Santorini (Greece). – 1997. – P.563-566.

А.В. Ніценко, Т.С. Хашан

Застосування комплексного безперервного вейвлет-перетворення Морле в обробці аудіосигналів

У статті запропоновано модифікований метод фазового вокодеру, що заснований на комплексному безперервному вейвлет-перетворенні Морле. Комбінація цього методу, лінійної інтерполяції та КІХ-фільтра дозволила розробити алгоритм зміни темпу сигналу без зміни висоти його тону.

A.V. Nitsenko, T.S. Khashan

Complex Continuous Morlet Wavelet Transform Application in Audio Signals Processing

In this article proposed a modified phase vocoder method, based on complex continuous Morlet wavelet transform. Combination this method with linear interpolation and FIR-filter allowed to develop algorithm of changing the audio tempo without altering the pitch.

Статья поступила в редакцию 29.07.2008.